Stochastic approximation properties in Banach spaces

V. P. Fonf[*], W. B. Johnson[†], G. Pisier[‡] and D. Preiss

August 16, 2002

Dedicated to A. Pełczyński on the occasion of his seventieth birthday.

1 Introduction

The paper deals with stochastic versions of the approximation property (AP) and the basis property (BP) of Banach spaces. Recall these concepts. One condition equivalent to saying that the Banach space X has the AP is that for any compact set $K \subset X$ there is a sequence of finite-dimensional (bounded) linear operators in X which pointwise converges to the identity on K. We say that a Banach space X has the BP if it has a (Schauder) basis. The stochastic versions are the following.

Given a Radon probability measure μ on a Banach space X, we say that X has the μ–approximation property (μ–AP, in short) provided there is a sequence $\{B_n\}$ of finite-dimensional operators on X so that $\|x - B_n x\| \to 0$ for μ almost every x in X. If only $B_n x \to x$ weakly for μ almost every x we say that X has the weak μ–AP.

We say that a separable X has the μ–basis property if there is an M-basis (that is, a sequence of vectors with dense linear span for which there is a sequence of biorthogonal functions which separates points) $\{x_n\}_{n=1}^\infty$ of X (with biorthogonal functionals $\{x_n^*\}_{n=1}^\infty$) for which

$$\mu\{x \in X : \ x = \sum x_n^*(x) x_n\} = 1;$$

$\{x_n\}_{n=1}^\infty$ is then called a μ–basis for X.

We say that a Banach space X has the stochastic AP (respectively, stochastic BP, for separable X) provided X has the μ–AP respectively, μ–BP) for every Radon

[*]Part of this research was done while a Visiting Professor at Texas A&M University.

[†]Supported in part by NSF DMS-9900185, DMS-0200690, Texas Advanced Research Program 010366-0033-20013, U.S.-Israel Binational Science Foundation, and a Weston Visiting Professorship at The Weizmann Institute of Science.

[‡]Supported in part by NSF DMS-9900185, DMS-0200690, and Texas Advanced Research Program 010366-0033-2001.

probability measure μ on X (see e.g. [T], where the stochastic AP is called the "measure approximation property"; for the stochastic BP see [H],[O]). The weak stochastic AP is defined analagously.

The stochastic AP is (formally) weaker than the AP, and the stochastic BP is (formally) weaker than the BP. It is well known (see [FJ]) that the AP is (really) weaker than the BP. It is even better known (see [LT1, Theorem 2.d.6]) that there are separable Banach spaces without the AP.

It is then natural to ask the following two questions:

(1) Is the stochastic approximation property really weaker than the stochastic basis property (as in the case of the AP and the BP)?

(2) Suppose that a Banach space X has the stochastic approximation property. Then must X have the approximation property?

Most of the results in this paper were discovered while working on these questions and variations of them.

In section 2 we prove that, although the AP and the BP are different, the stochastic AP and the stochastic BP are equivalent (see Theorem 2.1 that contains a stronger result and has an application to function theory).

The main result on question (2), Theorem 3.1 of section 3, is that the answer is affirmative if X has non trivial type. It follows, in particular, that not every Banach space has the stochastic approximation property, which answers a question asked by J. Rosinski and presented at a conference by S. Kwapien twenty-two years ago.

It is interesting that μ–stochastic properties for certain naturally occurring μ are of a different behavior. For example, each separable Banach (and even Fréchet) space has a μ–basis for every Gaussian probability μ (see [H] and [O] for a generalization).

In section 4 we discuss the weak stochastic AP and L_p versions of the AP. Recall that the AP and the weak AP are equivalent. We prove that the same holds for stochastic versions. We also observe that, for any fixed $1 \le p \le \infty$, the L_p approximation property is equivalent to the AP.

The last two sections are devoted to the problem of covering a measure support by an operator range. In section 5 we consider the following question. Assume that $K \subset X$ is a compact set in a (separable) Banach space X. Is it possible to find a probability μ on X such that for any linear operator $A : Z \to X$ from a Banach space Z into X with $\mu(A(Z)) = 1$ we have $A(Z) \supset K$? This question is a stochastic version of a problem investigated in [FJPS]. The answer is negative (see Theorem 5.1) and it shows that covering a compact set and covering a measure support by operator ranges are of a different nature.

In section 6 we consider the following situation. Suppose X and Y are separable Banach spaces such that for every Radon probability on X, there is an operator from Y to X whose range has probability one. The obvious way this can happen is for Y to be isomorphic to a quotient of X. In Corollary 6.3 we prove that this is the only way for this to occur when Y is $L_p(\equiv L_p[0,1])$, $1 < p < \infty$. This generalizes a theorem of Sato [Sa], who treated the case when Y is ℓ_2.

We use standard Banach space theory terminology, as may be found in [LT1],

[LT2].

2 Stochastic AP and stochastic BP are equivalent

The main result of this section is contained in the following

Theorem 2.1. *Let μ be a Radon probability measure on a separable Banach space X. Then the following assertions are equivalent:*
(i) X has the μ–approximation property.
(ii) X has a μ–basis.

Proof. Clearly, only $(i) \implies (ii)$ needs to be proved. We use in the proof some ideas from the proof of Theorem 2.1 in [FJPS]. Let $\{B_n\}_{n=1}^{\infty}$ be a sequence of finite dimensional operators on X which converges μ–almost everywhere to the identity operator on X. Put $Q = \{x \in X : \lim B_n x = x\}$. By using Egoroff's theorem find an increasing sequence $\{C_n\}$ of subsets of Q such that $\lim_n \mu(C_n) = 1$ and so that on each C_n the convergence $B_n x \to x$ is uniform. Take an index n_1 such that $\sup_{C_1} ||B_{n_1}x - x|| < 2^{-1}$. Next take $n_2 > n_1$ such that $\sup_{C_2} ||B_{n_2}x - x|| < 2^{-2}$, and so on. In this way we construct an increasing sequence of indices $\{n_k\}$ such that for each k, $\sup_{C_k} ||B_{n_k}x - x|| < 2^{-k}$. Put $C = \cup C_k$. Clearly, $\mu(C) = 1$. From the construction it follows that for each $k > 1$,

$$\sup_{C_k} ||(B_k - B_{k-1})x|| < 2^{-k+2}.$$

In particular, the series

$$\sum_{k=1}^{\infty} (B_k - B_{k-1})x, \; B_0 = 0$$

converges absolutely to x for any $x \in C$. For each k put $L_k = (B_k - B_{k-1})X$ and let Y be the Banach space of all sequences (y_k), $y_k \in L_k$, $k = 1, 2, ...$ such that the series $\sum y_k$ converges, with norm $||(y_k)|| = \sup_n || \sum_{k=1}^{n} y_k||$ (see Lemma 2.2 in [FJPS]). It is clear that Y has the bounded approximation property (and even a finite dimensional decomposition). Let $B : Y \to X$ be the summation operator, i.e. $B(y_k) = \sum y_k$.

Next define Z as the space of all $x \in X$ such that $x = \sum_{k=1}^{\infty} (B_k - B_{k-1})x$, $B_0 = 0$ with norm

$$|||x||| = \sup_n || \sum_{k=1}^{n} (B_k - B_{k-1})x||.$$

Let $J : Z \to Y$ be the natural isometry of Z into Y. Clearly, $M = J(Z)$ is a (closed) subspace of Y. Let $I : Z \to X$ be the natural embedding of Z into X. It is obvious that $I(Z) \supset C$ and $B|_M = IJ^{-1}$. Define a (probability) measure ν on Z by

$$\nu(G) = \mu(I(G))$$

where $G \subset Z$ is a Borelian subset of Z. Let $\{K_n\}$ be a sequence of compact sets in Z with $\nu(\cup K_n) = 1$. Put $a_n = \max_{x \in K_n} ||x||$, $n = 1, 2, ...$ and $K = \cup (a_n + n)^{-1} K_n$. Thus, K is compact in Z.

By Theorem 2.1 in [FJPS] there is a one to one compact operator $T_1 : R \to Y$ from a reflexive space R with basis into Y such that $T_1(B_R) \supset K$. Set $F = T_1(B_R)$ and put $L = \text{Ker } B$. It is not difficult to see that $L \cap M = \{0\}$.

Apply Lemma 2.6 in [FJPS] to find an automorphism $D : Y \to Y$ such that $L \cap D(F) \subset \{0\}$ and $D|_M = Id_M$. Finally, put $T = BDT_1$. A simple verification shows that T is 1-1 and $T(R) \supset C$.

By Lemma 2.10 in [FJPS] there are a 1-norming M-basis $\{x_i\}$ of X and a basis $\{y_i\}$ of R such that $\{Ty_i\} \subset \{x_i\}$. It is clear that $\{x_i\}$ satisfies condition (ii) of the theorem. □

Now we present an application of Theorem 2.1 to function theory.

Definition 2.2. *Let X be a subspace of $C[0,1]$ and μ a probability measure on the interval $[0,1]$. We say that an M-basis $\{x_i\}$ of X is a μ-quasi-basis of X if there is a subset $Q \subset [0,1]$ with $\mu(Q) = 1$ such that for any $x \in X$ the Fourier series $\sum_{i=1}^{\infty} x_i^*(x) x_i(t)$ converges to the function $x(t)$ for any $t \in Q$ ($\{x_i^*\}$ are the biorthogonal functionals for $\{x_i\}$).*

Remark 2.3. *If we substitute in the above definition convergence μ–a.e. by convergence everywhere (i.e. for any $t \in [0,1]$), then, at least for a subspace X which does not contain an isomorphic copy of c_0, we get that the notions of quasi-basis and basis coincide (see [F]).*

Corollary 2.4. *Assume that a subspace $X \subset C[0,1]$ has a separable dual X^* having the stochastic AP. Then X has a μ-quasi-basis for any probability measure μ on $[0,1]$.*

Proof. As usual, we consider $[0,1]$ as a subset of $B_{C^*[0,1]}$. Thus any probability measure μ on $[0,1]$ may be considered as a probability measure on $B_{C^*[0,1]}$ in the w^*-topology. Denote by $J : X \to C[0,1]$ the natural embedding of X into $C[0,1]$ and define a measure ν on X^* by $\nu(A) = \mu(J^{*1-}(A) \cap [0,1])$, where $A \subset X^*$ is a w^*-Borelian subset of X^*. Since X^* is separable it follows that the σ-algebra of w^*-Borelian sets coincides with σ-algebra of Borelian sets. Thus ν is a probability measure on X^*. Since X^* has the stochastic AP there is a $1-1$ operator $T : R \to X^*$ from a reflexive space R with basis into X^* such that $\nu(T(R)) = 1$ (see the proof of Theorem 2.1). Without loss of generality we may assume that $T(R)$ is dense in X^*. (Indeed, put $L = \text{cl}T(R)$ and let L_1 be a quasi-complement for L in X^*. Take any dense embedding $A : l_2 \to L_1$ and define $T_1 : (R \oplus l_2) \to X^*$ by $T_1(x + y) = Tx + Ty$, $x \in R$, $y \in l_2$. Finally, pass to R_1 and T_1.) Clearly, T is the adjoint operator for some operator $T_* : X \to R_*$ (R_* is the predual of R). Since T is one-to-one it follows that T_* has a dense range. Clearly, R_* has a basis $\{y_i\}$

with $y_i \in T_*(X),\ i = 1, 2, ...$ Denote by $\{y_i^*\}$ the biorthogonal functionals for $\{y_i\}$ and put $x_i = T_*^{-1} y_i$ and $x_i^* = T y_i^*,\ i = 1, 2, ...$ It is not difficult to see that $\{x_i\}$ is an M-basis of X with biorthogonal functionals $\{x_i^*\}$. Also it is clear that for any $f \in T(R)$, $f = \sum_{i=1}^{\infty} f(x_i) x_i^*$ (the series converges in the norm topology).

We show that $\{x_i\}$ is a μ-quasi-basis of X. Put $Q = J^{*-1}(T(R)) \cap [0,1]$. Clearly, $\mu(Q) = 1$. Take $x \in X,\ t \in Q,$ and check that $x(t) = \sum_{i=1}^{\infty} x_i^*(x) x_i(t)$. Since $J^*(t) \in T(R)$ it follows that $J^*(t) = \sum_{i=1}^{\infty} J^*(t)(x_i) x_i^*$, and hence

$$x(t) = J^*(t)(x) = \sum_{i=1}^{\infty} J^*(t)(x_i) x_i^*(x) = \sum_{i=1}^{\infty} x_i^*(x) x_i(t).$$

(Actually we used that the series $f = \sum_{i=1}^{\infty} f(x_i) x_i^*$ converges in the w^*-topology for any $f \in T(R)$.) □

Remark 2.5. *Corollary 2.4 shows, in particular, that a μ-quasi-basis need not be a basis (take any subspace $X \subset C[0,1]$ without a basis which has a separable dual with the AP; see [Sz]).*

3 Stochastic AP is equivalent to AP for spaces with non trivial type

Theorem 3.1. *Let X have a non trivial type and assume that X fails the AP. Then X fails the stochastic AP.*

To prove Theorem 3.1 we need two known lemmas.

Lemma 3.2. *Let $\{\theta_n\}_{n=1}^{\infty}$ be an independent, identically distributed sequence of symmetric one-stable (Cauchy) random variables on some probability space $(\Omega, \mathbb{P})$ and let let $\{y_n\} \subset Y$ be a sequence in a Banach space Y. Then for any $0 < p < 1$,*

$$\mathbb{E}\|\sum_k \theta_k y_k\|^p \geq \delta(\sum_k \|y_k\|)^p, \tag{3.1}$$

where $\delta > 0$ is a constant which depends only on p.

Proof. First lower estimate the left hand side of (3.1) by $\mathbb{E}\sup_k \|\theta_k y_k\|^p$

Next assume, without loss of generality, that $\mathbb{E}\sup_k \|\theta_k y_k\|^p = 1$. Then

$$1/2 > \mathbb{P}[\sup_k \|\theta_k y_k\|^p > 2] = 1 - \prod_k (1 - \mathbb{P}[\|\theta_k y_k\|^p > 2]),$$

so that

$$c2^{-1/p} \sum_k \|y_k\| \leq \sum_k \mathbb{P}[\|\theta_k y_k\|^p > 2] < \ln 2,$$

where the constant c satisfies $\mathbb{P}[|\theta_k| > t] \geq c/t$ for all $t > 1$. □

The following lemma is an easy consequence of Corollary 3.2 in [HJ1].

Lemma 3.3 ([HJ1]). *Let $\{\theta_n\}_{n=1}^{\infty}$ be an independent, identically distributed sequence of symmetric one-stable random variables on some probability space $(\Omega, \mathbb{P})$, $\{x_k\} \subset X$ a sequence in a Banach space X, and $\{\lambda_k\}$ a sequence of positive numbers with $\sum \lambda_k = 1$. Assume that $\{T_n\}$ is a sequence of finite rank operators on X such that*

$$\|\sum_k \lambda_k \theta_k (x_k - T_n x_k)\| \to 0$$

in measure as $n \to \infty$. Then

$$\mathbb{E}\|\sum_k \lambda_k \theta_k (x_k - T_n x_k)\|^p \to 0$$

for each $0 < p < 1$, as $n \to \infty$.

Proof of Theorem 3.1. Since X fails the AP it follows from the Grothendieck theorem [Gr], [LT1, Theorem 1.e.4], that there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in the unit ball of X, $\{x_n^*\}_{n=1}^{\infty}$ in the unit ball of X^* and a summable sequence $\{\lambda_n\}_{n=1}^{\infty}$ of positive numbers summing to one so that

$$\sum \lambda_n \langle x_n^*, x_n \rangle > \alpha > 0 \tag{3.2}$$

but

$$\sum \lambda_n \langle x_n^*, T x_n \rangle = 0 \tag{3.3}$$

for every finite rank operator T on X.

Let $\{\theta_n\}_{n=1}^{\infty}$ be an independent, identically distributed sequence of symmetric one-stable (Cauchy) random variables on some probability space $(\Omega, \mathbb{P})$. Consider the X valued random variable

$$\Phi = \sum_{k=1}^{\infty} \lambda_k \theta_k x_k$$

This series converges almost surely because X has type $r > 1$. Indeed, use the type inequality and then replace θ_k with $\theta_k 1_{[|\theta_k| < 1/\lambda_k]}$. Next use the fact that $P[\theta_k > t]$ is like $1/t$ for large t.

Thus Φ is a well-defined X-valued random variable and hence induces a Radon probability, say ν, on X. The claim is that X fails the ν-stochastic AP.

Assume to the contrary that there is a sequence $\{T_n\}$ of finite rank operators on X which converges ν-a.e. to the identity. Going back the random variable, this means that

$$\|\sum_k \lambda_k \theta_k (x_k - T_n x_k)\| \to 0$$

in measure, as $n \to \infty$. By Lemma 3.3 we get for each $0 < p < 1$ that

$$\mathbb{E}\|\sum_k \lambda_k \theta_k (x_k - T_n x_k)\|^p \to 0$$

as $n \to \infty$.

However by using Lemma 3.2 (for $y_k = \lambda_k(x_k - T_n x_k)$), (3.2) and (3.3) we get

$$\mathbb{E}\|\sum_k \theta_k(\lambda_k(x_k - T_n x_k))\|^p \geq \delta(\sum_k \lambda_k\|x_k - T_n x_k\|)^p \geq \delta(\sum \lambda_k|x_k^*(x_k - T_n x_k)|)^p > \delta\alpha^p, \tag{3.4}$$

a contradiction that completes the proof. □

Remark 3.4. *If X has the μ-stochastic AP for every compactly supported probability μ then X has the μ-stochastic AP for every Radon probability μ by Lemma 5.2. But it seems not very convenient to replace the 1-stable measure in the proof of Theorem 3.1 by some measure which has bounded support.*

Remark 3.5. *Theorem 3.1 answers negatively the following question of J. Rosinski (see Math. Nachr. 95 (1980), p.302, Problem 16; see also [T]): does every separable Banach space have the stochastic AP? Indeed, take any separable Banach space X with non trivial type that does not have the AP. By Theorem 3.1, X does not have the stochastic AP.*

In view of Theorem 3.1 the following conjecture seems to be natural

Conjecture. The properties AP and stochastic AP coincide.

4 Stochastic p-approximation property and weak stochastic AP

Definition 4.1. *Given a measure μ on a Banach space X we say that X has the $L_p(\mu)$–approximation property provided there is a sequence $\{S_n\}_{n=1}^{\infty}$ of continuous finite rank operators on X such that $\|S_n x - x\|$ tends to zero in $L_p(\mu)$.*

Definition 4.2. *Given a Banach space X and $0 \leq p \leq \infty$, say that X has the p–approximation property provided X has the $L_p(\mu)$–approximation property for every probability μ on X which has compact support.*

It is clear that if $0 \leq p < r \leq \infty$ and X has the r–approximation property then X has the p–approximation property. It is easy to see that X has the ∞–approximation property if and only if X has the approximation property. It follows from Lemma 5.2 that X has the 0–approximation property if and only if X has the stochastic approximation property. Only slightly less obvious is that for $0 < p < \infty$, the p–approximation property implies the $L_p(\mu)$–approximation property for all Radon measures which have bounded support.

Proposition 4.3. *If X has the p–approximation property, $0 < p < \infty$, and μ is a Radon probability on X which has bounded support, then X has the $L_p(\mu)$–approximation property.*

Proof. Assume, without loss of generality, that there are disjoint, totally bounded subsets K_n of B_X so that $\mu \cup K_n = 1$. Choose $1 \le a_n \uparrow \infty$ so that

$$\sum a_n^p \mu K_n < \infty \tag{4.5}$$

Define a measure ν on X by

$$\nu B = \sum a_n^p \mu \left((a_n B) \cap K_n \right).$$

Then ν is a finite Radon measure which is supported on the totally bounded set $\cup a_n^{-1} K_n$, hence X has the $L_p(\nu)$–approximation property. However, for any bounded linear operator T on X,

$$\begin{aligned} \int \|Tx\|^p d\mu(x) &= \sum \int_{K_n} \|Tx\|^p d\mu(x) &&= \sum \int_{a_n^{-1} K_n} \|Ty\|^p d\mu(a_n y) \\ &= \sum \int_{a_n^{-1} K_n} \|Ty\|^p d\nu(y) &&= \int \|Ty\|^p d\nu(y). \end{aligned}$$

Hence X has the $L_p(\mu)$–approximation property if and only if X has the $L_p(\nu)$–approximation property. □

The next theorem implies that for $1 \le p \le \infty$, the p–approximation property is equivalent to the approximation property. Note, however, that when X has non trivial type, Theorem 4.4 does not follow from Theorem 3.1 since it is clear that if μ is any discrete measure on a Banach space X then X has the μ–approximation property.

Theorem 4.4. *If X fails the approximation property then there is a compactly supported discrete probability μ on X so that X fails the $L_1(\mu)$–approximation property.*

Proof. By the Grothendieck theorem [Gr], [LT1, Theorem 1.e.4], there exists a norm null sequence $\{x_n\}_{n=1}^\infty$ in the unit ball of X, $\{x_n^*\}_{n=1}^\infty$ in the unit ball of X^* and a summable sequence $\{\lambda_n\}_{n=1}^\infty$ of positive numbers summing to one so that

$$\sum \lambda_n \langle x_n^*, x_n \rangle > \alpha > 0 \tag{4.6}$$

but

$$\sum \lambda_n \langle x_n^*, T x_n \rangle = 0 \tag{4.7}$$

for every finite rank operator T on X.

Let μ be the discrete probability which assigns mass λ_n to x_n for $1 \le n < \infty$. If T is a finite rank bounded linear operator on X then

$$\begin{aligned} \int \|(I-T)x\| d\mu(x) &= \sum \|(I-T)x_n\| \lambda_n \\ &\ge \sum \lambda_n \langle x_n^*, (I-T)x_n \rangle \\ &= \sum \lambda_n \langle x_n^*, x_n \rangle \ge \alpha. \end{aligned} \tag{4.8}$$

This shows that X fails the μ–AP. $\square$

For any fixed $p < 1$, we do not know whether the p–AP is equivalent to the AP for Banach spaces which have trivial type.

Now we pass to the weak stochastic AP. It is well-known that the weak AP is equivalent to the AP. The following theorem shows that the same is true for the stochastic version of the AP.

Theorem 4.5. *If X has the weak stochastic AP then X has the stochastic AP.*

Proof. Let μ be a separably supported probability on X and $\{S_n\}_{n=1}^{\infty}$ a sequence of finite rank operators on X for which there is subset A of X with $\mu(A) = 1$ and $S_n x \to x$ weakly for each x in A. For x in A set $M_x := \sup_n \|S_n x - x\| < \infty$. Given $\varepsilon > 0$, there is a compact subset B of A so that $\mu(B) > 1 - \varepsilon$ and $M := \sup_{x \in B} M_x < \infty$. Consider the compact Hausdorff space $K = B \times B_{X^*}$, where B is given the norm topology from X and B_{X^*} is given the weak* topology. For each n define f_n in $C(K)$ by

$$f_n(x, x^*) = \langle x^*, S_n x - x \rangle.$$

Then

$$\sup_n \|f_n\| := \sup_n \sup_{x \in B} \sup_{x^* \in B_{X^*}} \langle x^*, S_n x - x \rangle = M < \infty.$$

Also $\{f_n\}_{n=1}^{\infty}$ converges pointwise to zero on K, hence $f_n \to 0$ weakly in $C(K)$. Consequently, there is a convex combination $g := \sum\limits_{i=1}^{n} \lambda_i f_i$ for which $\sup_{t \in K} g(t) < \varepsilon$. But

$$\sup_{t \in K} g(t) = \sup_{x \in B} \|(\sum_{i=1}^{n} \lambda_i S_i) x - x\|.$$

Letting $\varepsilon = 2^{-1}, 2^{-2}, \ldots$, we get a sequence of finite rank operators on X which converges to the identity μ-a.e. $\square$

5 Covering a compact set by an operator range of full measure

The main result of this section is Theorem 5.1, which shows that in any infinite-dimensional Banach space X there is a compact set K so massive that for any Radon probability measure μ on X there is an operator range of probability one that does not contain K.

Theorem 5.1. *Let X be an infinite dimensional Banach space. Then there is a compact subset K of X so that if μ is any Radon probability on X then there is a bounded linear operator $A : Z \to X$ from a Banach space Z such that $\mu(A(Z)) = 1$ but K is not a subset of $A(Z)$.*

In the proof we use the following simple lemma:

Lemma 5.2. *Let μ be a Radon probability on a Banach space X. Then there is a Radon probability ν on X which is supported on a compact subset of B_X so that for any Borel subspace Y of X, $\mu Y = 1$ if and only if $\nu Y = 1$.*

Proof of Lemma 5.2. Take totally bounded, disjoint subsets K_n of X so that $\mu \cup K_n = 1$. Set $M_n := \sup_{x \in K_n} \|x\|$ and define measures ν_n on the Borel subsets of X by the formula $\nu_n A := \nu(nM_n A) \cap K_n$. It is routine to verify that the probability $\nu := \sum \nu_n$ has the desired properties. $\square$

The main technical result needed for the proof of Theorem 5.1 is Proposition 5.3.

Proposition 5.3. *Let μ be a Radon probability measure on a Banach space X. Then there is an operator from a Hilbert space into X with 2-summing adjoint whose range contains the subspace*

$$Y_0 := \cap\{Y \subset X : Y \text{ is an operator range and } \mu Y = 1\}. \tag{5.9}$$

First we show how to derive Theorem 5.1 from Proposition 5.3. If X is (isomorphic to) a Hilbert space, it suffices, by Proposition 5.3, to take for K any compact subset of X which is not contained in the range of Hilbert-Schmidt operator. For example, K can be $\{0\} \cup \{x_n\}_{n=1}^{\infty}$, where $\{x_n\}_{n=1}^{\infty}$ is an orthogonal sequence which converges to zero but $\sum \|x_n\|^2 = \infty$.

If X is not isomorphic to a Hilbert space, there there is even a compact subset K of X which is not contained in the range of *any* operator from a Hilbert space into X (see, for example, [FJPS]), so also in this case Theorem 5.1 follows from Proposition 5.3.

Proof of Proposition 5.3. Assume, without loss of generality, that X is separable. Let μ be a Radon probability on H. In view of Lemma 5.2, without loss of generality we may assume that $\mu B_H = 1$; in particular, $\int \|x\|^2 d\mu(x) < \infty$. It is well known that this implies that the mapping $S : X^* \to L_2(H, \mu)$ defined by $S(x^*)(x) = \langle x, x^* \rangle$ is 2-summing. Indeed, given $x_1^*, \ldots, x_n^*$ in H^* so that for all x in H we have $\sum_{i=1}^n |\langle x, x_i^* \rangle|^2 \le \|x\|^2$, we have that

$$\sum_{i=1}^n \|Sx_i^*\|^2 = \sum_{i=1}^n \int \sum_{i=1}^n |\langle x, x_i^* \rangle|^2 d(\mu x) \le \int \|x\|^2 d\mu(x).$$

Hence $\pi_2(S)^2 \le \int \|x\|^2 d\mu(x)$.

We can also assume, without loss of generality, that the operator S is one to one (for example, replace μ with the average of μ and a discrete measure whose support is dense in B_X). This is just for convenience of notation later.

Let us first treat the case when X is a Hilbert space. When μ is a Gaussian measure, it is well known that the range of S^* (called the reproducing kernel Hilbert space associated with the measure μ) is equal to Y_0. The point of Proposition 5.3 is that the proof of the inclusion $Y_0 \subset S^* L_2(H,\mu)$ does not require that μ be Gaussian. To see this, choose an orthonormal basis $\{e_n\}_{n=1}^\infty$ for $X = X^*$ so that $\{Se_n\}_{n=1}^\infty$ is orthogonal; say, $Se_n = \lambda_n f_n$ with $\lambda_n > 0$ and $\{f_n\}_{n=1}^\infty$ orthonormal. (Here we are using that S is one to one.) Then $S^* f_n = e_n$, $S^* f = 0$ if $f \perp \{f_n\}_{n=1}^\infty$, and $\sum_{n=1}^\infty \lambda_n^2 \le \pi_2(S)^2 < \infty$.

Fix any x_0 in x such that x_0 is not in the range of S^*; then $x_0 = \sum_{n=1}^\infty \alpha_n e_n$ with $\sum_{n=1}^\infty (\alpha_n/\lambda_n)^2 = \infty$. Define $g_n : X \to \mathbb{R}$ by

$$g_n(x) = \sum_{k=1}^n \frac{\langle x, e_k\rangle \alpha_k}{\lambda_k^2}.$$

Note that $g_n(x_0) = \sum_{k=1}^n (\alpha_k/\lambda_k)^2 \to \infty$. Compute

$$\|g_n\|^2 = \int g_n(x)^2 d\mu(x) = \sum_{k=1}^n (\alpha_k/\lambda_k)^2 \int \langle x, e_k\rangle^2 d\mu(x) = \sum_{k=1}^n (\alpha_k/\lambda_k)^2 .$$

Set $h_n = g_n / (\sum_{k=1}^n (\alpha_k/\lambda_k)^2)$. Then $\|h_n\|_{L_2(\mu)} = (\sum_{k=1}^n (\alpha_k/\lambda_k)^2)^{-1/2} \to 0$ and $h_n(x_0) = 1$ for all n. Thus some subsequence $\{h_{n_k}\}_{k=1}^\infty$ of $\{h_n\}_{n=1}^\infty$ converges to zero μ-almost everywhere. Let $X_0 = \{x \in X : h_{n_k}(x) \to 0\}$. Then $\mu X_0 = 1$ and x_0 is not in X_0.

To prove that X_0 is an operator range introduce on X_0 a new norm as follows

$$|||x||| = \max\{\|x\|,\ \max\{|h_{n_k}(x)| :\ k = 1,2,...\}\},\ \ x \in M.$$

A standard verification shows that $(X_0, |||.|||)$ is a Banach space (call it Z) and that the natural embedding $A : Z \to X$ is a bounded linear operator, which completes the proof of Proposition 5.3 in the case X is a Hilbert space.

Joel Zinn pointed out to us that the general case in Proposition 5.3 follows easily from the Hilbert space case (thereby rendering silly a slightly involved argument of the authors). So now assume that X is a separable Banach space and μ is a Radon probability on X so that the support of μ is B_X. Recall that the operator $S : X^* \to L_2(H,\mu)$ defined by $S(x^*)(x) = \langle x, x^*\rangle$ is one to one and S^* is 2-summing. Note also that S is weak* to weak sequentially continuous by the bounded convergence theorem, and hence $S^* L_2(X,\mu) \subset X$. To prove (5.9) by reducing to the Hilbert space case, we let J be a one to one bounded linear operator with dense range from X into a Hilbert space H and let ν be the image measure on H of μ; that is, $\nu E = \mu J^{-1} E$ for every Borel subset E of H. Let $\phi_J : L_2(H,\nu) \to L_2(X,\mu)$, defined by $\phi_j f = fJ$, be the natural surjective isometry and consider the commutative diagram

$$\begin{tikzcd} L_2(H,\nu) \arrow[drr, "\phi_J"] & & \\ H^* \arrow[u, "S_H"] \arrow[r, "J^*"] & X^* \arrow[r, "S"] & L_2(X,\mu) \end{tikzcd}$$

where S_H is defined by $S_H(x^*)(x) = \langle x, x^* \rangle$. Keeping in mind that $S^* L_2(X,\mu) \subset X$ we dualize to get the commutative diagram

$$\begin{tikzcd} L_2(H,\nu)^* \arrow[d, "S_H^*"'] & & \\ H & X \arrow[l, "J"'] & L_2(X,\mu)^* \arrow[l, "S^*"'] \arrow[ull, "\phi_J^*"'] \end{tikzcd}$$

Since ν is the image of μ under J, (5.9) follows from the Hilbert space case of Proposition 5.3. Indeed, if $x_0 \in X \sim S^* L_2(X,\mu)^*$ then $Jx_0 \in H \sim S_H^* L_2(H,\nu)^*$, so there is an operator range $H_0 \subset H$ for which $\nu H = 1$ but $Jx_0 \notin H_0$. It is not hard to check that $X_0 \equiv J^{-1} H_0$ is an operator range, and of course $\mu X_0 = 1$ while $x_0 \notin X_0$. □

Remark 5.4. *The operator range X_0 constructed in the proof of Theorem 5.1 is an $F_{\sigma\delta}$ set.*

6 Covering a measure support by the image of a Banach space.

Say that the pair (Y, X) of separable Banach spaces satisfies $(*)$ provided that for any Radon probability μ on X there is a bounded linear operator T from Y into X so that $\mu(TY) = 1$ (or is positive; it is the same property).

We will denote $Y_T = Y/\ker T$ and $\widehat{T}\colon\ Y_T \to X$ will be the operator canonically associated to T when T is an operator from Y into X.

Lemma 6.1. *Assume that (Y, X) satisfies $(*)$. Then for any $\varepsilon > 0$ there is a constant $\varphi(\varepsilon)$ such that for any Radon probability measure μ on B_X and for any $\varepsilon > 0$, there is $T\colon\ Y \to X$ with $\|T\| \le 1$ such that*

$$\mu(\{x \in X \mid \|\widehat{T}^{-1} x\|_{Y_T} \ge \varphi(\varepsilon)\|x\|\}) \le \varepsilon.$$

Proof. We may clearly assume that $\mu(\{0\}) = 0$. Then, replacing μ by its image under the map $x \to x/\|x\|$, we see that it suffices to prove this statement assuming μ is supported on the unit sphere of X. We will then proceed by contradiction. If the conclusion fails, then there exists $\varepsilon > 0$ such that for any $\varphi > 0$ we can find a random variable

$$g_\varphi\colon\ \Omega \to X$$

(on some standard probability space $(\Omega, \mathcal{A}, \mathbb{P})$) such that $\|g_\varphi(\omega)\|_X = 1$ but such that for any $T\colon\ Y \to X$ with $\|T\| \le 1$ we have

$$\mathbb{P}\{\|\widehat{T}^{-1} g_\varphi\|_{Y_T} \ge \varphi\} > \varepsilon.$$

Let us choose $\varphi = n^3$. For simplicity of notation we set $f_n = g_{n^3}$. We have then $\|f_n(\cdot)\| = 1$ and

$$\mathbb{P}\{\|\widehat{T}^{-1}(f_n)\| \geq n^3\} > \varepsilon \tag{6.10}$$

for any T with $\|T\| \leq 1$ and any n.

We may as well assume that the random variables (f_n) are mutually independent and also that we have a sequence (ε_n), itself independent of (f_n), of i.i.d. Bernoulli random variables on Ω such that $P\{\varepsilon_n = \pm 1\} = 1/2$. Then the series $S = \sum_1^\infty n^{-2}\varepsilon_n f_n$ (with partial sums $S_n = \sum_1^n k^{-2}\varepsilon_k f_k$) obviously converges a.s. in X. By our assumption $(*)$ there is $T\colon\ Y \to X$ with $\|T\| \leq 1$ such that $S(\omega) \in TY = \widehat{T}Y_T$ for a.e. ω.

Observe that, since $S_n \pm (S - S_n)$ and S have the same distribution, we also have $S_n - (S - S_n) \in TY$ a.s., and hence $S_n \in TY$ a.s. for any n. A fortiori, $f_n \in TY$ a.s. Consider now the random variable $\omega \to \widehat{T}^{-1}(S(\omega)) \in Y_T$. Note that its distribution must be Radon (recall Y is separable) and moreover we have obviously

$$\langle \xi, \widehat{T}^{-1}(S_n)\rangle \to \langle \xi, \widehat{T}^{-1}(S)\rangle$$

for any ξ in $\widehat{T}^*(X^*)$. But since $\widehat{T}$ is injective, $\widehat{T}^*(X^*)$ is a w^*-dense w^*-separable subspace of Y_T^*. Therefore by a variant of the Ito–Nisio theorem (see, for example, Theoem 6.2 in [HJ2] we must have "automatically"

$$\|\widehat{T}^{-1}(S_n) - \widehat{T}^{-1}(S)\|_{Y_T} \to 0 \quad \text{a.s.}$$

This implies $\sup\limits_n n^{-2}\|\widehat{T}^{-1}(f_n)\| < \infty$ a.s. hence (Borel–Cantelli) for some constant K

$$\sum_n P(\|\widehat{T}^{-1}(f_n)\| > Kn^2) < \infty.$$

But since $Kn^2 \leq n^3$ for infinitely many n, this contradicts (6.10). This contradiction completes the proof. □

Theorem 6.2. *Assume that (Y, X) satisfies $(*)$ and that $Y \oplus Y$ is isomorphic to a quotient of Y. Then for any $1 \leq p < \infty$ there is a constant C such that the following holds: for any n and any $x_1, \ldots, x_n$ in X there is $T\colon\ Y \to X$ with $\|T\| \leq 1$ and $y_1, \ldots, y_n$ in Y such that $x_i = Ty_i$ for all $i = 1, \ldots, n$ and*

$$\left(\sum \|y_i\|^p\right)^{1/p} \leq C\left(\sum \|x_i\|^p\right)^{1/p}.$$

Proof. Fix n. Let C_n be the smallest constant such that the preceding property holds when restricted to n-tuples $(x_1, \ldots, x_n)$. Clearly $C_n < \infty$. To prove the preceding statement it suffices to show that $\sup_n C_n < \infty$. By Lemma 6.1 applied to the probability $\mu = \left(\sum \|x_i\|^p\right)^{-1} \sum \|x_i\|^p \delta_{x_i/\|x_i\|}$ (and by the definition of the quotient norm in Y_T), for any $\varepsilon > 0$ there are $T\colon\ Y \to X$ with $\|T\| \leq 1$, a subset $A \subset \{1, \ldots, n\}$ and elements $\{y_i \mid i \in A\}$ in Y such that

$$\forall i \in A \qquad Ty_i = x_i \quad \text{and} \quad \|y_i\| < \varphi(\varepsilon)\|x_i\| \forall i \in A$$

and such that

$$\sum_{i \notin A} \|x_i\|^p \leq \varepsilon \sum \|x_i\|^p.$$

By the definition of C_n, there is $S\colon\ Y \to X$ with $\|S\| \leq 1$ and $\{y_i \mid i \notin A\}$ in Y such that $x_i = Sy_i$, $i \notin A$, and

$$\begin{aligned} \left(\textstyle\sum_{i \notin A} \|y_i\|^p\right)^{1/p} &\leq C_n \left(\textstyle\sum_{i \notin A} \|x_i\|^p\right)^{1/p} \\ &\leq C_n \varepsilon^{1/p} \left(\textstyle\sum \|x_i\|^p\right)^{1/p}. \end{aligned}$$

Let $v\colon\ Y \oplus_1 Y \to X$ be defined by $v(y_1, y_2) = (Ty_1 + Sy_2)/2$. Let $\widetilde{y}_i \in Y \oplus Y$ be defined by $\widetilde{y}_i = 2(y_i, 0)$ if $i \in A$ and $\widetilde{y}_i = 2(0, y_i)$ otherwise. We have then $v\widetilde{y}_i = x_i$ and

$$\left(\sum \|\widetilde{y}_i\|^p\right)^{1/p} \leq (2\varphi(\varepsilon) + 2C_n\varepsilon^{1/p}) \left(\sum \|x_i\|^p\right)^{1/p}.$$

Since $Y \oplus_1 Y$ is isomorphic to a quotient of Y, we may replace $Y \oplus_1 Y$ by Y, but then we find

$$\left(\sum \|\widetilde{y}_i\|^p\right)^{1/p} \leq K(2\varphi(\varepsilon) + 2C_n\varepsilon^{1/p}) \left(\sum \|x_i\|^p\right)^{1/p}$$

for some constant K depending only on Y. Thus we conclude

$$C_n \leq K(2\varphi(\varepsilon) + 2C_n\varepsilon^{1/p}),$$

and it suffices to choose $\varepsilon = \varepsilon_0$ such that e.g. $2K\varepsilon_0^{1/p} = 1/2$ to obtain finally

$$C_n \leq K2\varphi(\varepsilon_0) + C_n/2.$$

Hence

$$C_n \leq 4K\varphi(\varepsilon_0),$$

which completes the proof that $\sup_n C_n < \infty$. □

Corollary 6.3. *Let $Y = L_p([0,1])$ with $1 < p < \infty$. Then (Y, X) satisfies $(*)$ if and only if X is isomorphic to a quotient of Y.*

Proof. If X is a quotient of Y then $(*)$ trivially holds. Conversely, assume $(*)$. It obviously suffices to show that X is isomorphic to a quotient of an abstract L_p-space. The latter spaces, which we will call QL_p-spaces, admit the following very nice characterization ([LP]; see also [K2] for a more complete picture): X is C-isomorphic to a QL_p-space iff for any n, any quotient Q of ℓ_p^n and any linear mapping $u\colon\ \ell_p^n \to Q$ we have

$$\|u \otimes I_X\colon\ \ell_p^n(X) \to Q(x)\| \leq C\|u\|.$$

Here if $Q = \ell_p^n/S$ with $S \subset \ell_p^n$ the space $Q(X)$ is defined as the quotient $\ell_p^n(X)/S(X)$ where $S(X) = \{(x_1, \ldots, x_n) \mid (\xi(x_1), \ldots, \xi(x_n)) \in S\ \forall \xi \in X^*\}$ or equivalently in tensor product notation $S(X) \simeq S \otimes X$. Thus it suffices to show that the preceding characteristic property of QL_p-spaces is implied by the conclusion of Theorem 6.2 in the case $Y = L_p$.

Let $u\colon\ \ell_p^n \to Q$ be as above with $\|u\| \le 1$. Fix $x = (x_i) \in \ell_p^n(X)$. Let $y = (y_i) \in \ell_p^n(Y)$ and $T\colon\ Y \to X$ be as in Theorem 6.2 such that $\|T\| \le 1$, $x_i = Ty_i$ and $\|y\| \le C\|x\|$. We have then (since $Y = L_p$)

$$\|u \otimes I_Y\colon\ \ell_p^n(Y) \to Q(Y)\| \le 1$$

hence since

$$(u \otimes I_X)(x) = (u \otimes T)(y) = (I \otimes T)(u \otimes I_Y)(y)$$

we find

$$\begin{aligned} \|(u \otimes I_X)(x)\|_{Q(X)} &\le \|T\| \|(u \otimes I)(y)\|_{Q(Y)} \\ &\le \|y\| \le C\|x\|. \end{aligned}$$

Thus we conclude that X is C-isomorphic to a quotient of L_p. □

In particular, when $p = 2$, we obtain a new proof of the following result due to H. Sato [S]:

Corollary 6.4. *(ℓ_2, X) satisfies $(*)$ if and only if X is isomorphic to ℓ_2.*

Using Roberto Hernandez' generalization [He] of Kwapień's results from [K2] we can obtain by an immediate modification of the proof of Corollary 6.3 the following statement valid for a general Y.

Corollary 6.5. *Assume (Y, X) satisfies $(*)$. Then for any $1 \le p < \infty$ the inclusion $X \to X^{**}$ can be factorized through a quotient of an ultrapower of $L_p(Y)$.*

We do not know whether the conclusion of Corollary 6.5 can be strengthened to "the inclusion $X \to X^{**}$ can be factorized through a quotient of an ultrapower of Y".

The following easy lemma is surely folklore.

Lemma 6.6. *Let X be any infinite dimensional Banach space and let K be a compact subset of a Hilbert space H. Then there is a compact operator T from X into H so that TB_X contains K.*

Hint for proof. Using Dvoretzky's theorem and standard techniques for constructing basic sequences, one sees that X has a quotient space, Y, on which there is a uniformly bounded sequence $\{P_n\}_{n=1}^\infty$ of projections so that $P_nP_m = 0$ for $n \ne m$ and P_nX is 2–isomorphic to ℓ_2^n. (If X is separable, by using the technique for constructing weak* basic sequences one can even guarantee that $\{P_nX\}_{n=1}^\infty$ is a finite dimensional decomposition for Y.) □

The next proposition, which in some sense goes oppositely to Corollary 6.4, is an immediate consequence of Lemma 6.6.

Proposition 6.7. *Let μ be a Radon probability measure on a Hilbert space H and let X be an infinite dimensional Banach space. Then there is a bounded linear operator from X into H so that $\mu(TX) = 1$.*

References

[E] P. Enflo, "A counterexample to the approximation property in Banach spaces," *Acta Math.* 130 (1973), 309-317.

[FJ] T. Figiel and W. B. Johnson, "The approximation property does not imply the bounded approximation property," *Proc. Amer. Math. Soc.* 41 (1973). 197-200.

[F] V. P. Fonf, "Weakly extremal properties of Banach spaces," *Mat. Zametki*, 45 (1989), 83-92; English transl. in Math. Notes Acad. Sci. USSR 45 (1989), 488-494.

[FJPS] V. P. Fonf, W. B. Johnson, A. M. Plichko, and V. V. Shevchyk, "Covering a compact set by an operator range of a Banach space with basis," Trans. Amer. Math. Soc. (to appear).

[Gr] A. Grothendieck, "Produits tensoriels topoligiques et spaces nucleaires," *Memoirs Amer. Math. Soc.* 16 (1955).

[H] W. Herer, "Stochastic bases in Fréchet spaces", *Demonstratio Mathematica* 14 (1981), 719-724.

[He] R. Hernandez, "Espaces L^p, factorisation et produits tensoriels dans les espaces de Banach," (French; English summary) [L^p spaces, factorization and tensor products in Banach spaces] *C. R. Acad. Sci. Paris Ser. I Math.* 296 (1983), no. 9, 385-388.

[HJ1] J. Hoffman-Jørgensen, "Sums of independent Banach space valued random variables", *Aarhus Universitet, Matematisk Institut,* Preprint Series 1972/73, No. 15.

[HJ2] J. Hoffmann–Jørgensen, "Sums of independent Banach space valued random variables," *Studia Math.* 52 (1974), 159-186.

[K1] S. Kwapień, "Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients," *Studia Math.* 44 (1972), 583-595.

[K2] S. Kwapień, "On operators factorizable through L_p space," Actes du Colloque d'Analyse Fonctionnelle de Bordeaux (Univ. de Bordeaux) (1971), 215-225. Bull. Soc. Math. France, Mem. No. 31-32, Soc. Math. France, Paris, 1972.

[LP] J. Lindenstrauss and A. Pełczyński, "Absolutely summing operators in L_p-spaces and their applications," *Studia Math.* 29 (1968), 275-326.

[LT1] J. Lindenstrauss and L. Tzafriri, "Classical Banach spaces I: sequence spaces," Ergebnisse der Mathematik und ihrer Grenzgebiete 92, Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[LT2] J. Lindenstrauss and L. Tzafriri, "Classical Banach spaces II: function spaces," Ergebnisse der Mathematik und ihrer Grenzgebiete 97, Springer-Verlag, Berlin-Heidelberg-New York, 1979.

[O] Y. Okazaki, "Stochastic basis in Fréchet space", *Math. Ann.* 274 (1986), 379-383.

[Sa] H. Sato, "Hilbertian support of a probability measure on a Banach space," *Probability in Banach spaces II,* Springer Lecture Notes in Math. 709 (1979), 195-205.

[Sz] S. J. Szarek, "A Banach space without a basis which has the bounded approximation property," *Acta Math.* 159 (1987), 81-98.

[T] V. I. Tarieladze, "On the topological description of characteristic functionals in infinite dimensional spaces, *Probability in Banach spaces,* 217-232, Birkhauser, J. Hoffman-Jorgensen et al. (eds.), 1994.

W. B. Johnson
Department of Mathematics
Texas A&M University
College Station, TX 77843 U.S.A.
johnson@math.tamu.edu

V. Fonf
Department of Mathematics
Ben-Gurion University of the Negev
P. O. Box 653 Beer-Sheva 84105 Israel
fonf@math.bgu.ac.il

G. Pisier
Department of Mathematics
Texas A&M University
College Station, TX 77843 U.S.A.
pisier@math.tamu.edu
and
Equipe d'Analyse, Case 186, 75252
Université Paris VI
Paris, Cedex 05, France
gip@ccr.jussieu.fr

D. Preiss
Department of Mathematics
University College London
London, Great Britain
dp@math.ucl.ac.uk